The Impact of Cultural Diversity on Economic Growth in China

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation
with honors research distinction in Economics in the
undergraduate colleges of The Ohio State University

by

Zhen Ye

The Ohio State University
April 2013

Project Advisor: Professor Belton M. Fleisher, Department of Economics

Abstract

The paper examines the impact of cultural diversity on the economic growth in China. Based on provincial data of 29 Chinese provinces from 2000 to 2010, I find that ethnic diversity has negative impact on total output and linguistic diversity reduces economic growth. These effects are hypothesized to come from corruptions, conflicts between groups and higher communication costs. I conclude that publicizing cultures of different ethnic groups and popularizing the standard spoken Chinese would enhance economic performance and promote economic growth.

1. Introduction

Several empirical studies in growth economics explores the influence of diversity on the process of economic growth. Empirical evidence, mostly from cross-country studies, suggests that diversity has an impact on economic growth. For instance, Easterly and Levine (1997) show that ethnic diversity generates conflicts which lead to poor economic performance. In this paper, I report research on the role of cultural diversity defined as ethnic diversity and linguistic diversity on economic growth in China. I believe the results have important implications for an better understanding of economic growth in China in general as well as policies concerning education in China.

The paper is structured as follows. In section 2 I explore the motivation and review of the literature on cultural diversity and its major results. Section 3 lays out the models and variables. In section 4 I provide a description of data used to test the impact of ethnic diversity and linguistic diversity on economic growth in China. Section 5 reports estimated results. In section 6 I provide a robustness check. Section 7 concludes and provides policy recommendations.

2. **Motivation& Literature Review**

Provinces of China have experienced rather different growth paths under economic reform. While some of these provinces' outstanding performances are typical of the overall growth of China, some of the others have fallen behind (average per capita GDP growth shown in Figure 2). Several studies investigate the factors affecting regional economic inequality in China. Most of these studies focus on the impact of economic factors. For example, Demurger (2001) argues that appropriate infrastructure has positive impact on economic growth. Fleisher, Li and Zhao (2007) show that Foreign Direct Investment (FDI) and human capital have affected total factor productivity (TFP) positively and significantly and therefore have a great impact on economic growth. Demurger (2001) argues that agriculture has a negative impact on economic growth as agricultural provinces have fewer opportunity to have productivity growth. The influence of cultural diversity is largely overlooked while several studies, such as Easterly and Levine (1997) and Alesina et al (2003) show that ethnic and linguistic diversity are strongly correlated with economic growth.

Previous studies have provided mixed results concerning diversity's impact on economic growth. According to Easterly and Levine (1997), public policy choices in an ethnically diverse society are not economically optimal due to conflicts of preferences. Triesman (2000) and Glaeser and Saks (2006) argue that ethnic diversity reduces institutional quality and lead to corruption, which have negative impact on economic growth as shown by Mo (2001). Both Delhey and Newton (2005) and Alesina and La Ferrara (2002) find a negative relationship between diversity and trust, which is shown by Knack and Keefer (1997) to influence the incentive for innovation, accumulation of human capital and therefore economic growth. A rather recent study by Dincer and Wang(2011) on ethnic diversity and economic growth in China also find that ethnic diversity affects economic growth negatively through raising corruption in China. On the other hand, using U.S data, Ottaviano and Peri (2005) and Ratna, Grafton and Kompas (2009) find that linguistic diversity contributes positively

to economic growth. According to Ottaviano and Peri(2005), different skills from different cultures contribute to the productivity of native workers. One reason for the uncertainty of cultural diversity's impact is that the impact of diversity is, according to Collier(2000), only counterproductive in the presence of limited political rights.

Indeed, China's wide regional disparities in economic growth pattern, level of ethnic and linguistic diversity provide a very important and useful episode to analyze the effect of cultural diversity on economic growth. China has totally 56 officially recognized ethnic groups. However, the level of ethnic diversity is far from identical throughout China. According to Dincer and Wang(2011), inland China is about four times as ethnically diverse as its coastal counterpart, mostly because the five ethnic autonomous provinces are all located in the western China. China is also very linguistically diverse as a whole as the Chinese language is divided into a number of subgroups by linguists. More over, it is usually not mutually intelligible between these linguistic groups ("Varieties of Chinese", 2013). What makes it more interesting is that, in contrary to the case of ethnic diversity, coastal provinces are much more linguistically diverse compared to their inland counterpart due to the large number of migrants from other provinces in the past two decades. Based on my own estimate, for instance, the level of linguistic diversity in Shanghai is approximately five times higher than the one in Jilin in both 2000 and 2005; Guangdong province is about 30 times as linguistically diverse as its neigh Guangxi province. I hypothesize that cultural diversity has played an essential causal role on economic growth.

3. Methodology

To test the impact of cultural diversity, I specify and estimate two different models (level vs. growth):

3.1 Estimation method: level

I first estimate the impact of cultural diversity on province-level economic performance in 2000 and 2005. My framework is based on the widely-used Cobb-Douglas production function:

$$Y_{i,t} = A_{i,t} K_{i,t}^{\beta_k} L_{i,t}^{\beta_L} \varepsilon$$

where Y is aggregate output, A is total factor productivity, K is capital stock, L is labor, ε is a disturbance term, for province i=1, 2,.. n and for year t=1,2. β_K, β_L are the output elasticities of capital and labor respectively. To test the impact of explanatory variables on output per capita, I divide both sides of the production equation by L and take natural logarithm to make the equation linear:

$$\ln y_{i,t} = \ln(\frac{Y_{i,t}}{L_{i,t}}) = \ln A_{i,t} + \beta_K \ln K_{i,t} + (\beta_L - 1) \ln L_{i,t} \varepsilon$$

where y is real output per capita. I follow the framework that has been widely applied to assume a linear correlation between other explanatory variables, including ethnic and linguistic diversity, and ln$A$.

To measure the level of cultural diversity, I follow Easterly and Levine (1997), Alesina et al. (2003) and most other literature on diversity to use a fractionalization index defined as

$$ethnic_i / language_i = 1 - \sum_{j=1}^{N} n_{ij}^2$$

where *ethnic* and *language* are ethnic and linguistic fractionalization respectively and $n_{i,j}$ is the population share of ethnic/linguistic group j in province i. N is the total number of ethnic/linguistic groups. Therefore the fractionalization index is the probability that two random individuals in a province belong to two different

ethnic/linguistic groups. The value of index increases as the number of groups increases and reaches a maximum value close to 1 when no pairs of individuals belong to a same group. It reaches minimum of 0 when all people in the province are of the same ethnic or linguistic group. The estimated levels of ethnic and linguistic diversity are shown in Figure 3 and Figure 4, respectively.

On the basis of existing literature on economic growth in China, I include a set of other control variables to minimize the omitted-variable bias. I follow Fleisher and Chen (1997), Demurger (2001) to control for infrastructure, which is shown to be positively correlated with economic growth. I measure infrastructure as the total share of transportation post and telecommunications in gross domestic product (GDP). I follow Demurger (2001) to include share of agriculture in GDP, who argues that agricultural provinces have slower productivity growth. Following Fleisher, Li and Zhao (2007), I control for Foreign Direct Investment(FDI), which embodies foreign technology. I also control for two dummy variables *coast* including Beijing and *minor* since coast provinces have geographical advantage for trading and coastal provinces, Beijing and the five ethnic autonomous provinces have received preferential policies from the central government. Finally I control for the intersection term *language*FDI* because according to Florida (2002), heterogeneity will have a positive impact on economic performance provided the gains from trade exceed the costs of trade.

Therefore my full model for testing the effect of cultural diversity on province-level economic performance is as follows:

$$\ln y_{i,t} = c + \beta_K \ln K_{i,t} + \beta_L \ln L_{i,t} + \alpha\, ethnic_{i,t} + \delta\, language_{i,t} + \sum_{k=1}^{6} \beta_k X_{k_{i,t}} + u_{i,t} \qquad (1)$$

where ethnic and language are ethnic and linguistic diversity respectively and X_k are *infrastructure, agriculture, FDI, coast, minor* and the intersection term of *language&FDI.* A concern about model above is endogeneity of linguistic diversity, because it is likely that provinces attract immigrants due to their prosperity, which

leads to higher level to linguistic diversity, Thus we use instrumental variable estimation to correct for this possibility. I choose real disposable income of urban households per capital annual as the instrumental variable for linguistic diversity because it is commonly accepted that the relatively high income of urban attracts workers to immigrate and is not an explanatory variable of economic performance. I follow Dincer and Wang(2011) to estimate the model by seemingly unrelated regression (SUR), as it allows serial correlation.

3.2 Estimation method: growth

I estimate the effect of cultural diversity on economic growth from 2000-2005 and 2005-2010 by the following equation:

$$\frac{1}{T-t}\ln(\frac{y_{i,T}}{y_{i,t}}) = c + \lambda \ln y_{i,t} + \beta_1 physical\_capital_{i,t} + \beta_2 human\_capital_{i,t} + \alpha\, ethnic_{i,t} + \delta\, language_{i,t} + \sum_{k=1}^{6} \gamma_k X_{k_{i,t}} + u_{i,t} \quad (2)$$

where $\frac{1}{T-t}\ln\frac{y_{i,T}}{y_{i,t}}$ is the average economic growth rate of province i from initial year t to T. $\ln y_{i,t}$ captures the transitional dynamics when the economies are not in their steady states (qtd, in Ratna, Grafton&Kompas, 2009). Following Dincer and Wang(2011), I estimate human capital (*human_capital)* and physical capital (*physical_capital)* by share of population with at least a senior high school degree and share of gross fixed capital formation in GDP, respectively. Other explanatory variables(the Xs) are the same as equation (1). Equation (2) is a framework widely used in previous relevant literature, such as Ratna, Grafton& Kompas(2009) and Dincer and Wang (2011). Same as equation(1), I estimate equation(2) by seemingly unrelated regression.

4. Data

My data of GDP, GDP per capita in equation (2), infrastructure, FDI, labor (total employment), agricultural product, gross fixed capital formation are from various years of China Statistical Yearbook provided by China Data Online. My data of total population, population of each ethnic group and population with at least a senior high school degree are from China's Population Census (2000, 2005(1%)) provided by China Data Online. My data of GDP per capita is calculated as the quotient of GDP and labor, as shown in sec. 3.1, since equation (1) is based on Cobb-Douglas Production Function. Nominal data are deflated by GDP deflator from the World Bank Data. My data of capital stock is from Wu(2009).

I exclude Tibet and Inner Mongolia and have all other 29 provinces in the sample, because a number of different languages are spoken in each of Tibet and Inner Mongolia, most of which are not mutually intelligible and far different from Chinese. Estimation of level of linguistic diversity in these two provinces is likely to be strongly biased. I choose t=2000, 2005 in both of equation (1) and (2) and T=2005,2010 in equation (2). The selection of time is based on the availability of demographic data needed for linguistic diversity.

I estimate the ethnic diversity in each province with the share of population of groups Han, Mongol, Hui, Tibetan, Uyghur, Miao, Yi, Zhuang, Bouyei, Korean, Manchu, Dong, Yao, Bai, Tujia, Hani, Kazakh, Dai, Li and collection of all other ethnic groups because the data of population from each other individual group in each province is not available in the 2005 population census. The estimated level of ethnic diversity is considerably close to Dincer& Wang's estimate (2011), given the figures provided.

To estimate the share of population from each linguistic group, I sum up shares of population registered in each of provinces belonging to the same linguistic groups, based on the Map of Chinese Dialects (Figure 1) in China from www.llmap.org and

the list of varieties of Chinese ("Varieties of Chinese", 2013). I assume people speak the same variety of Chinese and have the same accents if they are registered in the same provinces or belong to the same linguistic groups. Share of Hakka group is combined with Yue (Cantonese) group due to lack of data, and that most people from Hakka group are assumed to be capable of Cantonese since Hakka group is mostly within Guangdong province, where Yue (Cantonese) is most popular. Pinghua group and Hakka dialect are also omitted due to lack of data and relative small area the languages cover.

The measure of both ethnic and linguistic diversity does not take foreign population into account.

5. Empirical Results

5.1 Results of level estimation

The results of the effect of cultural diversity on province-level economics performance are given in Table 1.

The estimated coefficient of ethnic diversity is negative and significant at 10% level, which agrees with previous result that ethnic diversity is negatively correlated with economic performance. One possible explanation is that ethnic diversity leads to corruption, according to Dincer and Wang (2011). They test for this hypothesis by assuming a positive and significant relationship between level of corruption in a province and the size of state owned enterprises (SOEs) and running a regression of size of SOES on ethnic diversity and other control variables and find a positive result.

On the other hand, both coefficients of linguistic diversity and the intersection term of linguistic diversity and FDI are not statistically significant, which is consistent with results in Ratna et al. (2009) based on US data. It suggests that no evidence

supports a relationship between linguistic diversity and province-level economics performance.

The estimated coefficients of other control variables have the hypothesized signs with the exception of *infrastructure*. I hypothesize that this is due to the relative small sample size and that different provinces focus on different aspects of infrastructure. The high standard error of estimated coefficients of *FDI* and the dummy variable *coast* are also likely suffering from the problem of small sample size.

5.2 Results of growth estimation

The estimated results of the impact of cultural diversity on economic growth rate are given in Table 2.

Both the estimated coefficients of *language* and the intersection term of *language* and *FDI* are statistically significant at 1% and 10% levels respectively. The negative sign of coefficient of linguistic diversity is consistent with the findings in Alesina et al. (2003) that linguistic fragmentation has a negative impact on per capita growth and the hypothesis that linguistic diversity makes communication more costly and therefore inhibits the dissemination of knowledge and technology. The positive sign of *language*FDI* agrees with the findings of Florida (2002) that linguistic diversity can contribute positively to economic growth given the gains from trade exceed the costs of trade.

The estimated coefficient of ethnic diversity is, however, not significant. The coefficients of other control variables mostly have the hypothesized signs and again with the exception of *infrastructure.* The explanations of the negative sign of *infrastructure* and the high standard error of some other control variables are the same as in Sec. 5.1.

6. Robustness checks

In this section, I present the regression results to check the effect of linguistic diversity and the intersection term of linguistic diversity and FDI with another measure of the diversity level. While varieties of Chinese are often not mutually intelligible, the entire population in China, with the possible exception of Tibet and Inner Mongolia, are taught in standard mandarin as required by law. Therefore it is reasonable to allow for the possibility that linguistic barriers to communication across social groups are due to different identifiable accents rather than differences between varieties of Chinese. I assume people from different provinces have distinguishable accents, which creates a barrier to communication between people, and re-estimate the level of linguistic diversity in each province by

$$language_i = 1 - \sum_{j=1}^{30} n_{j,i}^2$$

where $n_{j,i}$ is the share of population registered in province j in province i. I assume people from Sichuan and Chongqing have the same accents. The data source is China Population Census (2000, 2005(1%)) provided by China Data Online. The estimated results of level and growth are given in Table 3 and 4 respectively. The newly measured linguistic diversity is *language2*. As shown in table 3 and 4, the estimated results shown in table 3 and 4 are generally consistent with earlier results in terms of coefficients' signs and significance levels. The estimated impact of the intersection term of linguistic diversity and FDI is slightly weaker.

In summary, my estimated impact of the intersection term of linguistic diversity and FDI is sensitive to the measure of linguistic diversity. However, the estimated impact of ethnic diversity on level and of linguistic diversity on growth are robust.

7. Conclusion and policy recommendations

Dramatic dispersion between China's provinces in level economic performance and rate of growth has been observed since the inception of economic reform in 1978. I investigate the impact of cultural diversity, which has been usually overlooked, on province-level economic performance and growth. I hypothesize there is causal relationship between regional pattern of economic growth and cultural diversity.

The following estimated results are robust to alternative measure of diversity. First, ethnic diversity has a negative impact on province-level economic performance. According to previous study on ethnic diversity, such as Alesina et al. (2003) and Dincer and Wang (2011), ethnic diversity affects economic performance by generating inter-group conflicts and corruption. Second, linguistic diversity has a negative impact on economic growth. Previous findings such as Alesina et al. (2003) and Grafton et al. (2007) suggest that the negative impact is due to higher cost of communication and the consequent inhibition of dissemination of productivity enhancing knowledge. My results with linguistic diversity measured by share of population from different linguistic groups support results of Florida (2002) that high level of linguistic diversity benefits economic growth in economically vibrant provinces, where gains from trade exceed the costs.

The results of my research shows that it is important to publicize cultures of different ethnic groups and popularize the standard spoken Chinese commonly known as Beijing mandarin, both for reasons of reducing unnecessary conflicts between cultural groups and economic efficiency. Indeed, China's central government has passed laws popularizing mandarin since 1955. However, the law started to become most effective only after mid 2000s, when almost the entire rural population are guaranteed the nine years of education through elementary school to middle school funded by government.

References

Alesina,A.,Devleeschauwer, A.,Easterly, W.,Kurlat, S.,& Wacziarg, R. (2003) Fracctionalization. Journal of Economic Growth, 8 (2), 155-194.

Alesina, A. & La Ferrara, E.(2002). Who trusts others? Journal of Public Economics, 85, 207-234.

Alesina, A.,& La Ferrara, E. (2005). Ethnic diversity and economic performance. Journal of Economic Literature, 43 (3), 762-800.

China data online : Zhongguo shu ju zai xian(2013) : All China Marketing Research Co., Ltd. (access 2012, 2013)

Collier,P.,(2000). Ethnicity,politics and economic performance. Economics and Politics, 12(3), 225-245.

Delhey,J.&Newton,K (2005). Predicting cross-national levels of social trust: global pattern of Nordic exceptionalism? European Sociological Review, 21, 311-327.

Demurger, S. (2001). Infrastructure development and economic growth: An explanation for regional disparities in China? Journal of Comparative Economics, 29,95-117.

Dincer, O., Wang, F (2011). Ethnic diversity and economic growth in China. Journal of Economic Policy Reform, 14, 1-10.

Easterly, W.,&Levine, R.(1997). Africa's growth tragedy: Policies and Ethnic Divisions. Quarterly Journal of Economics, 112, 1203-1250.

Fleisher, B.M.,&Chen, J. (1997). The coast0noncoast income gap, productivity and regional economic policy in China. Journal of Comparative Economics, 25,220-236.

Fleisher, B. M. & Li, H. & Zhao, M. Q., (2007). Human Capital, Economic Growth, and Regional Inequality in China, IZA Discussion Papers 2703, Institute for the Study of Labor (IZA).

Florida,R.,(2002). Bohemia and economic geography. Journal of Economic Geography. 2(1), 55-71.

Glaeser, E.,Saks,R.(2006). Corruption in America. Journal of Public Economics, 90,1053-1072.

Grafton, R.Q., Kompas, T.,Owen, D.P., (2007). Bridging the barriers: knowledge connections, productivity and capital accumulation. Journal of Productivity Analysis 27 (3), 219 – 231.

Knack,S.& Keefer, P(1997). Does social capital have an economic payoff? A cross-country investigation. Quarterly Journal of Economics, 112, 1251-1288.

Map of Beijing Mandarin (2012). http://www.llmap.org. (access March, 2013)

Mo, P. H., (2001). Corruption and Economic Growth. Journal of Comparative Economics 29, 66-79.

Ottaviano, G.I.P., Peri, G., (2005). Cities and culture. Journal of Urban Economics 58, 304–337.

Ratna, N. N., Grafton, R. Q. & Kompas, T., (2009). Is diversity bad for economic growth? Evidence from state-level data in the US. The Journal of Socio-Economics 38, 859-870.

Treisman, D.(2000). The causes of corruption: a cross-national study. Journal of Public Economics, 76,399-457.

Varieties of Chinese. (2013). In Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Varieties_of_Chinese&oldid=549819116 (access March 2013)

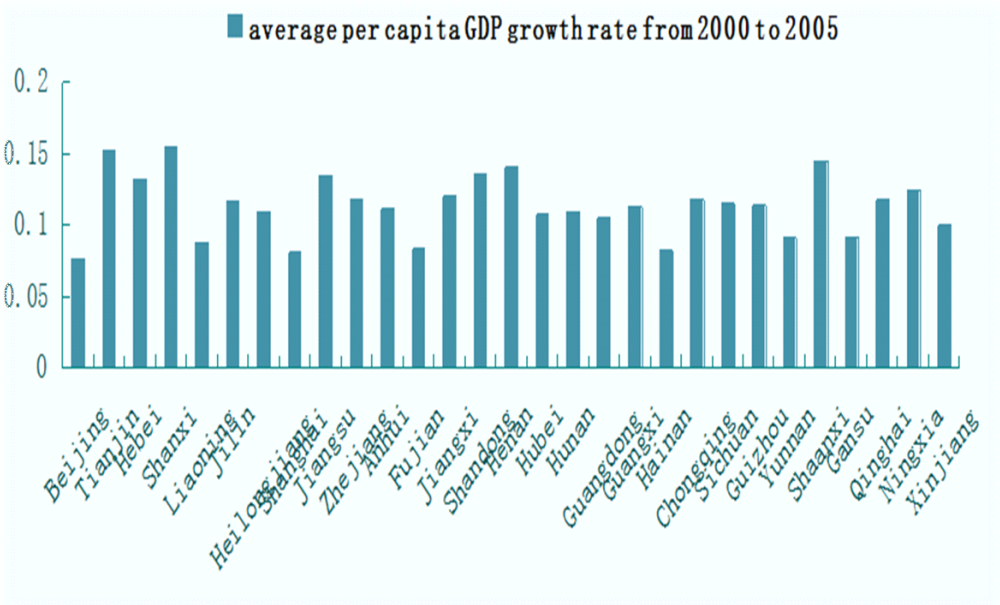
World Bank (2013). The World Bank Group. (access 2013)

Wu, Y. (2009). China's capital stock series by region and sector. Discussion paper. Business School, University of Western Australia.

Figure 1. Map of varieties of Chinese

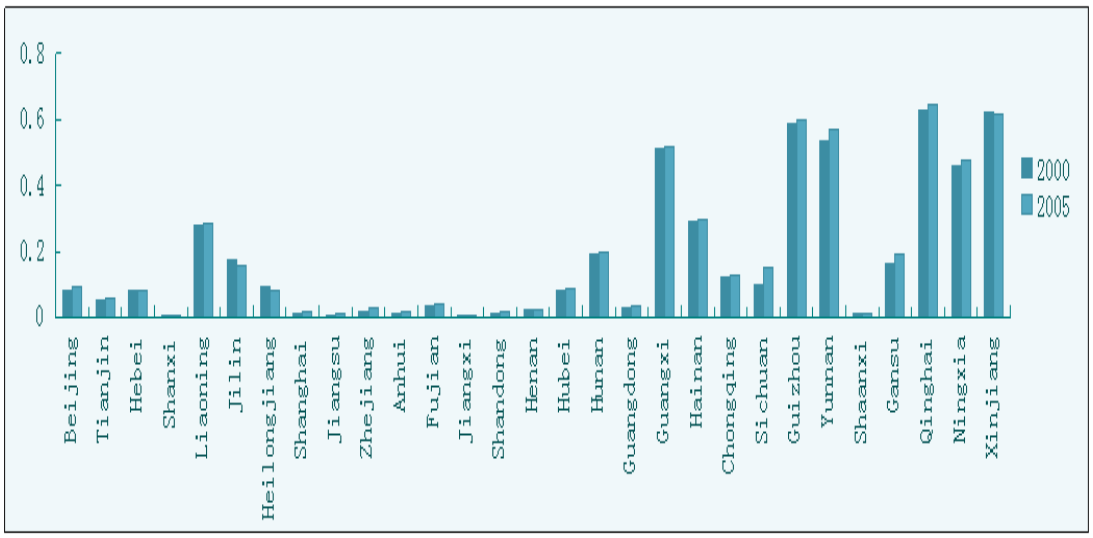
Source: www.llmap.org

Figure 2



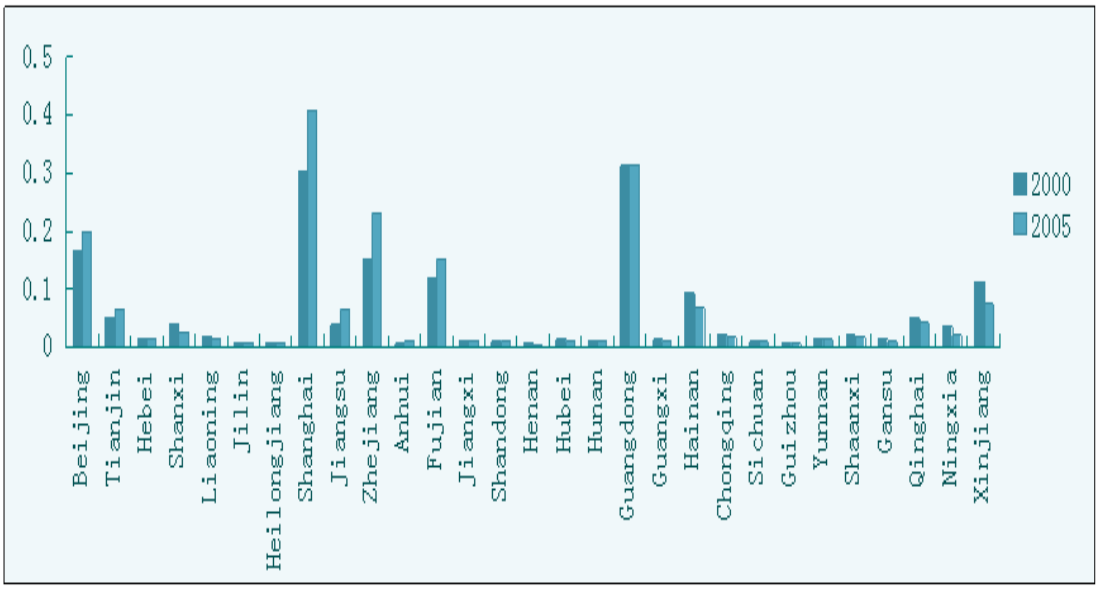
Source: China Statistics Yearbook

Figure 3. Level of each province's ethnic diversity in 2000 & 2005



Data Source: China Population Census

Figure 4. Level of each province's linguistic diversity in 2000 & 2005



Data Source: China Population Census

Table 1. Equation (1). Effect of diversity on province-level economic performance

Dependent Variable: log(real GDP per capita)		
Variable	Coefficient	Prob.
C	0.655	0.3048
AGRI	-1.143	0.0624
LOG(CAPITAL)	0.495	0.0002
COAST	0.149	0.1840
ETHNIC	-0.520	0.0558
FDI	0.00165	0.2291
INFRA	-1.042	0.3524
LOG(LABOR)	-0.513	0.0000
LANGUAGE	1.577	0.3259
MINOR	0.141	0.3450
LANGUAGE*FDI	-0.01083	0.4638

Table 2. Equation (2). Effect of cultural diversity on growth

Dependent Variable: per capita GDP growth		
Variable	Coefficient	Prob.
C	0.0525	0.3122
AGRI	-0.0735	0.0941
COAST	0.00639	0.3265
ETHNIC	-0.0189	0.2035
FDI	-0.000260	0.1320
HUMAN_CAPITAL	-0.1262	0.0096
INFRA	0.118	0.4791
LOG(INIPERGDP)	0.00816	0.1228
LANGUAGE	-0.180	0.0052
MINOR	0.000450	0.9572
PHYSICAL_CAPITAL	0.10012	0.0023
LANGUAGE*FDI	0.001619	0.0635

Table 3. Equation (1). Effect of cultural diversity on province-level economic performance with different measure of linguistic diversity

Dependent Variable: log(real per capita GDP)		
Variable	Coefficient	Prob.
C	0.528	0.3958
AGRI	-1.10024	0.0685
LOG(CAPITAL)	0.534	0.0000
COAST	0.1559	0.1451
ETHNIC	-0.501	0.0499
FDI	0.00102	0.4471
INFRA	-0.906	0.3976
LOG(LABOR)	-0.539	0.0000
LANGUAGE2	0.739	0.4775
MINOR	0.146	0.3018
LANGUAGE2*FDI	-0.00225	0.8292

Table 4 Equation (2) effect of cultural diversity on economic growth with different measure of linguistic diversity

Dependent Variable: Per capita GDP growth		
Variable	Coefficient	Prob.
C	0.0509	0.5097
AGRI	-0.2197	0.0018
COAST	0.0126	0.2104
ETHNIC	-0.0128	0.5821
FDI	-0.000522	0.0408
HUMAN_CAPITAL	-0.166	0.0556
INFRA	0.382	0.0976
LOG(INIPERGDP)	0.0124	0.1187
LANGUAGE2	-0.215	0.0268
MINOR	-0.00349	0.7914
PHYSICAL_CAPITAL	0.0855	0.0511
LANGUAGE2*FDI	0.00168	0.1664